Investigating Cyclic Peptides Inhibiting CD2-CD58 Interactions through Molecular Dynamics and Molecular Docking Methods

Leherte, Laurence; Petit, Axel; Jacquemin, Denis; Vercauteren, Daniel; Adèle, Laurent

Published in:
Journal of computer-aided molecular design

DOI:
10.1007/s10822-018-0172-4

Publication date:
2018

Document Version
Peer reviewed version

Investigating Cyclic Peptides Inhibiting CD2-CD58 Interactions through Molecular Dynamics and Molecular Docking Methods

Laurence Leherte[1*], Axel Petit[1], Denis Jacquemin[2,3], Daniel P. Vercauteren[1], Adèle D. Laurent[2]

[1] Laboratoire de Physico-Chimie Informatique, Unité de Chimie Physique Théorique et Structurale, Department of Chemistry, NAmur MEdicine & Drug Innovation Center (NAMEDIC), Namur Institute of Structured Matter (NISM), University of Namur, Rue de Bruxelles 61, B-5000 Namur (Belgium)

[2] University of Nantes, CEISAM UMR CNRS 6230 UFR sciences et techniques, 2 Rue de la Houssinière BP 92208, F-44322 Nantes Cedex 03 (France)

[3] Institut Universitaire de France, 103 Bd St Michael, F-75005 Paris Cedex 5 (France)

ORCID: L. Leherte 0000-0001-8468-5462
D. Jacquemin 0000-0002-4217-0708
D. Vercauteren 0000-0002-6164-8551
A. Laurent 0000-0001-9553-9014

Acknowledgments

The authors thank the reviewers for their comments that helped to improve the manuscript, as well as Dr. Jose Ceron-Carrasco and Horacio Pérez Sánchez for numerous discussions. Frédéric Wautelet and Laurent Demelenne are gratefully acknowledged for program installation and maintenance. The research used resources of the 'Plateforme Technologique de Calcul Intensif (PTCI)' (http://www.ptci.unamur.be) located at the University of Namur, Belgium, which is supported by the F.R.S.-FNRS convention 2.5020.11. The PTCI is member of the 'Consortium des Équipements de Calcul Intensif (CÉCI)' (http://www.ceci-hpc.be). This research used as well French resources of (1) the GENCI-CINES/IDRIS (Grants A0020805117) and (2) CCIPL (Centre de Calcul Intensif des Pays de Loire). The authors also thank the Interuniversity Attraction Pole program n° 7/05: 'Functional supramolecular systems' initiated by the Belgian Science Policy Office.

This work is supported by the Wallonie-Bruxelles International (WBI) and the Belgian National Foundation for Scientific Research (FNRS), by the French Ministry of Foreign and European Affairs, and by the Ministry of Higher Education and Research, in the framework of the Hubert Curien partnerships (PHC Tournesol #40638PL).

[*] Corresponding author: Email. laurence.leherte@unamur.be, Tel. +32-81-724560, Fax. +32-81-725466

Abstract

The CD2-CD58 protein-protein interaction is known to favor the recognition of antigen presenting cells by T cells. The structural, energetics, and dynamical properties of three known cyclic CD58 ligands, named P6, P7, and RTD-c, are studied through Molecular Dynamics (MD) simulations and molecular docking calculations. The ligands are built so as to mimic the C and F β-strands of protein CD2, connected via turn inducers. The MD analyses focus on the location of the ligands with respect to the experimental binding site and on the direct and water-mediated hydrogen bonds (Hbonds) they form with CD58. Ligand P6, with a sequence close to the experimental β-strands of CD2, presents characteristics that explain its higher experimental affinity, e.g., the lower mobility and flexibility at the CD58 surface, and the larger number and occurrence frequency of ligand-CD58 Hbonds. For the two other ligands, the structural modifications lead to changes in the binding pattern with CD58 and its dynamics. In parallel, a large set of molecular docking calculations, carried out with various search spaces and docking algorithms, are compared to provide a consensus view of the preferred ligand binding modes. The analysis of the ligand side chain locations yields results that are consistent with the CD2-CD58 crystal structure and suggest various binding modes of the experimentally identified hot spot of the ligands, i.e., Tyr86. P6 is shown to form a number of contacts that are also present in the experimental CD2-CD58 structure.

Keywords CD58, cyclic peptides, ligands, molecular docking, molecular dynamics, inhibitor

Introduction

Protein-protein interactions (PPI) allow, e.g., cells to communicate with each other and form a huge complex network, called interactome, which significantly contributes to the biological processes that are carried out in organisms. Up to 650,000 PPI networks have been reported so far, making them highly attractive targets to design new drugs compared to single proteins [1]. Such interactions are involved in the majority of human life regulation and deregulation that might lead to a disease state, e.g., they control signal transduction, immune response, and transcription [2]. Therefore, targeting PPI with organic molecules such as peptides is a promising approach in modern drug discovery.

A detailed knowledge of the interaction surfaces of the involved proteins, their energetics and dynamics, are helpful to understand the regulatory mechanisms of biochemical pathways aiming to modulate or block these pathways for therapeutic purposes. The contact surfaces of protein-protein interfaces can be relatively large (ranging from 1,000 to 4,000 $Å^2$) and flat, inducing a real challenge for the design of inhibitors. In this paper, we focus on a rather small interface (1,200 $Å^2$) involving the cell adhesion molecule CD2 found on the surface of T cells (and natural killer cells), and its counter-receptor CD58 present on antigen-presenting cells (APC) [3]. CD2 plays an essential role in the cell-cell interactions and signal transductions through interactions with its CD58 counterparts in humans. Blocking molecules like CD58 results in obstructing the T cell receptor/APC interactions and in preventing the primary immune response. Particularly, the CD2-CD58 protein-protein interface is reported as crucial in inflammatory and autoimmune diseases such as rheumatoid arthritis [4]. One of the most promising approach to inhibit CD2-CD58 PPI is to develop peptides mimicking the CD2 interaction site with CD58 [5,6].

The interfacial domain of CD2-CD58 is planar and mainly involves electrostatic interactions that ensure a high specificity between both partners [7]. CD2 is reported to transiently bind to its receptor CD58 with a low affinity, i.e., K_d = 0.4 μM [8]. At the molecular level, a hot spot has been identified and involves Tyr86 of CD2 and Lys34 of CD58 [9]. The electrostatic potential at the CD58 surface is mainly negative, due to glutamate and aspartate residues, while CD2 mostly presents positive potential values at its interface with CD58, due to the presence of arginine and lysine residues. Particularly, the CD58 binding domain of CD2 consists of β-strands C and F with charged residues. Peptides mimicking the CD2 β-strands C and F have already been synthesized to specifically block the PPI [5,6], while paying a careful attention to their solubility and geometrical constraints of the β-strands motif. Particularly, three cyclic peptide ligands reported in literature are characterized by low IC_{50} values, i.e., P6 [10], P7 [11], and RTD-c [12], with values of 6.9 ± 0.4 [6,12], 11.1 ± 3.8 [6], and 27 ± 15 nM [12], respectively (Fig. 1).

Previous MD simulations of the CD2-CD58 complex corroborate that electrostatics govern the PPI, showing that salt bridges determine the protein-protein binding strength [13]. Bayas et al. carried out steered MD simulations of the solvated CD2-CD58 system using various velocity and force schemes at room temperature. They characterized the progressive breaking of the salt bridges occurring under protein-protein detachment, and later confirmed the results with experimental force measurements on various mutants [14]. Based on a tool

specifically dedicated to the analysis of MD trajectories, MDcons, Abdel-Azeim et al. studied the contacts occurring in the solvated low affinity CD2-CD58 complex at 300 K compared to a high affinity complex [15]. They related the lower affinity of CD2-CD58 to its larger interface flexibility. More recently, through MD simulations of the solvated CD2-CD58 complex, Wang et al. determined that the binding of CD2 with CD58 is governed by the topology of a set of three hot spots in CD2, i.e., Lys41, Lys51, and Gly90. Additionally, they showed that glycosylation of CD2 induces a shorter protein-protein distance with a stable and reduced number of interfacial water molecules [16].

The identification of important residues in the CD58 receptor and its ligands is also achieved using molecular docking approaches. In particular, Jining et al. used the Autodock program to dock linear and cyclic 12-mer and hexapeptide ligands onto CD58. No direct correlation was found between the best poses and their biological activity; however, an agreement was pointed out with antibody inhibition assay [17]. The ligands appeared to interact with Lys29, Asp33, and Lys34 of CD58, which are listed as being important in the CD2-CD58 binding. Docking results of two ligands named P6 and RTD-c were obtained with Autodock by Gokhale et al. [10] and Sable et al. [12], respectively. From results involving ligand P6, the authors suggested that CD58 residues such as Lys29, Lys30, Asp33, Lys34, and Phe46, are important in the binding of CD2 to CD58, while for RTD-c, residues Lys29, Lys32, Lys34, Glu42, and Arg44 of CD58 play a significant role. Such works clearly indicate that both cyclic peptides are not located in the same area. Additionally, P6 is reported to bind to the CD2-binding site of CD58 [10]. As reviewed by Lawson et al. [18], small peptide ligands built from residues involved in crucial contacts within a protein-protein interface might no longer reproduce the adequate interactions with the experimental binding site of the receptor, even if the peptide is structurally constrained. Nevertheless, these small ligands still could likely bind efficiently to the receptor, at different binding locations [18].

In the present work, we characterize, at the molecular level, the structure, energetics, and dynamics of ligand-CD58 complexes using MD simulations and molecular docking calculations. Three ligand are considered, i.e., P6, P7, and RTD-c (Fig. 1). Trajectories are analyzed to characterize the effect of water molecules and hydrogen bonds on the positioning of the ligand on the CD58 surface. Differently, molecular docking calculations are carried out with constrained search spaces starting from structures generated by MD simulations. To deal with the flat CD58 binding surface, a consensus analysis of multiple docking results obtained

from various software and calculation conditions is achieved. It allows the identification of favored ligand-CD58 contacts and interactions.

Materials and computational Methods

Ligand structures

As mentioned in the Introduction, three peptide ligands were selected from literature based on their low IC_{50} value, i.e., P6, P7, and RTD-c, with values of 6.9 ± 0.4, 11.1 ± 3.8, and 27± 15 nM, respectively. All of them adopt a cyclic structure and are designed to mimic the interacting region of the CD2 protein, i.e., residues 31 to 34 and 84 to 87 of the C and F antiparallel β-strands (Fig. 1 and Table 1). RTD-c involves two additional residues that are not contained in P6 and P7, i.e., Glu36 and Lys82. Those two CD2 residues do not form neither a hydrogen bond (Hbond) nor a salt bridge with CD58 according to the contacts identified with the server PDBePISA [19] applied to the CD2-CD58 crystal structure (PDB code 1QA9) [20]. However, a contact between Glu36 of CD2 and Glu78 of CD58 at a distance shorter than 0.5 nm is observed in the 1QA9 crystal structure, as reported through the Protmap2D program [21]. Ligands P6 and P7 contain a (D)-proline (L)-proline motif and a dibenzofuran derivative (DBF), respectively, to induce the β-turn between the two β-strand segments, while RTD-c is cyclized by two alanine-glycine turns and is constrained by three disulfide bridges. Ligands P6 and RTD-c involve 10 and 18 amino acid residues, respectively, while P7 is composed of only seven residues including the DBF moiety (Fig. 1).

The cartesian coordinates of structures P6 and P7 were obtained using the *fragment builder* option of PyMol, in antiparallel β-sheet mode [22]. Regarding the RTD-c inhibitor, an NMR structure of the Rhesus theta defensin (RTD-1) in an antiparallel β-strand conformation was retrieved from the Protein Data Bank (PDB ID 2LYF). The required mutations were applied using the PyMol *Mutation* wizard. A reference peptide ligand, named REF further in the text, was also built using the initial 3D crystal structure data of the eight residues located in the C and F strands of protein CD2, i.e., residues Asp31 to Lys34 and Ser84 to Asp87, respectively (Fig. 1). The two β-strands were linked with CH_2 groups, selected to avoid the occurrence of polar interactions with CD58. Further details regarding the building and preparation (atomic charges) of the final ligand structures are reported in Online Resource 1.

Molecular Dynamics simulations

In all calculations, the chain B of the PDB structure 1QA9 was taken to build the receptor CD58. It is composed of 95 amino acid residues, with a net charge of -3 $|e^-|$, while ligands bear a total net charge of -2 $|e^-|$. Five K^+ ions were thus taken into account to neutralize the electric charge of the ligand-CD58 complexes. The coordinate and topology files of the receptor CD58 were automatically generated using the *pdb2gmx* tool of Gromacs, which allows the assignment of H atoms and the histidine protonation state, Hisε. Due to their cyclic structure, the creation of the corresponding files for the ligands using the current Gromacs tools was not straightforward. The full procedure used in the present work is reported in Online Resource 2. The ligands were finally placed at a position close to the two β-strands of the CD2 partner in its crystallographic orientation (Online Resource 2, Fig. OR2.1).

MD simulations of the whole systems were run in vacuum (see Online Resource 3 for data analysis) and in water using the Gromacs4.5.5 program package [23,24] with the Amber99sb FF [25]. To strongly reduce the calculation time of the solvated systems, the hybrid TIP3P/SIRAH water FF was used [26-28]. In solution, the systems were submitted to short MD runs of 10 to 50 ps to progressively increase the temperature from 50 to 300 K, a temperature value that was previously considered in other MD works regarding CD2-CD58 systems [13-16]. Next, each system was equilibrated during 60 ns MD at 300 K and 1 bar until convergence of the RMSD profile and confinement of the ligand in a single area of the CD58 surface (Online Resources 3 and 4). In that last case, the distance between the center of mass (com) of the ligand and the com of all CD58 binding residues identified with PDBePISA was considered, regardless of the ligand orientation. A final production run of 100 ns (50 x 10^6 steps) in vacuum and 300 ns in water (150 x 10^6 steps) was performed for the evaluation of the structural, energetics, and dynamical properties of each system.

Molecular docking calculations

As for the MD simulations, the receptor CD58 is taken from the crystallographic coordinate file (PDB ID 1QA9) and the protonation stage of His is set to Hisε. The atomic coordinates of ligands P6, P7, and RTD-c, were obtained from the central structure generated through the clustering of the MD trajectory conformations of the vacuum complexes (Online Resource 3,

Fig. OR3.3). Additionally, a fourth ligand named REF was built from the data reported in the PDB structure 1QA9 to serve as a reference structure (Fig. 1).

As specified in the Introduction, the CD58 does not present any well-defined binding cavity shape. Docking is thus expected to lead to many possible solutions unless specific and strong intermolecular interactions could systematically favor the expected binding mode(s). To detect such interactions, several docking programs, scoring functions, and search spaces were considered, and a consensus analysis was carried out over the best docking poses. Three well-known molecular software programs were selected to perform the docking calculations of the ligands at the surface of protein CD58, i.e., Autodock 4.2 [29], Smina [30], a computer program based on Autodock Vina [31], Gold [32]. Details regarding these programs are reported in Online Resource 5. Docking calculations were also achieved using several protein-protein rigid docking web servers like ZDOCK [33], FRODOCK [34], and CLUSPRO [35], as well as the flexible protein-protein docking program HADDOCK [36].

To define the search space in ligand-docking calculations, several grid centers were considered, i.e., (i) the NZ atom of Lys34 of CD58, (ii) the geometrical center of residues 31 to 34 and 84 to 87 of CD2, (iii) the geometrical center of the CD58 site constituted by a limited number of residues, i.e., 27, 29, 32 to 34, 37, 44, 46, and 78, and (iv) the geometrical center of the 18 interface residues identified by PDBePISA [19]. The four models are respectively named *m_lys34*, *m_CFstrands*, *m_smallsite*, *m_largesite* (Online Resource 5). In model (iii), i.e., *m_smallsite*, the nine CD58 residues are all located within a distance of 0.5 nm from the amino acid residues 31 to 34 and 84 to 87 of CD2.

In Autodock and Smina docking calculations, the receptor and ligand coordinate files were prepared following the protocols reported by Forli et al. [37]. All molecular structure elements were kept rigid except for the side chains of the ligand. To define the search space, grid parameters such as the size and a center were required. The grid size was determined using the eBoxSize Tool [38]. All search space parameters are reported in Online Resource 5, Table OR5.3.

With Gold, all molecules are flexible. To define the search space, several cavity definitions were used, i.e., either the selected center is *m_CFstrands* or *m_lys34*, and the cavity radius is 10, 12, or 15 Å (named R10, R12, and R15 further in the text). Two additional calculations were carried out using a cavity defined by the residues of model *m_smallsite* and *m_largesite*, as described earlier.

On the whole, 20 molecular docking calculations were carried out for each ligand.

Results and discussion

Molecular Dynamics calculations

The 3D atomic coordinates of the receptor CD58 were taken from the crystallographic coordinate file (PDB ID 1QA9) and the initial configuration of the ligand-CD58 complexes were obtained by manually placing the ligands at a position close to the two β-strands of the CD2 partner in its crystallographic orientation (Online Resource 2, Fig. OR2.1).

Ligand location. The final MD trajectories of each ligand-receptor system were first analyzed to inspect the location of the ligands at the surface of CD58. In vacuum, a superimposition of the ligand structures obtained during the 100 ns MD simulations shows that even if the poses remain at locations close to residues of the CD58 binding site (Online Resource 3, Fig. OR3.4), none of them coincide perfectly with the expected orientation found in the crystal structure of the CD2-CD58 complex (PDB ID 1QA9) (Fig. 2). In water, the ligands, especially P7, are more distantly placed from the crystal structure binding site (Fig. 3 and Online Resource 4, Fig. OR4.2). Indeed, the average distance between the com of P7 and the com of the experimental CD58 binding residues is close to 2 nm, to be compared to values of about 1.36 and 1.24 nm for P6 and RTD-c, respectively (Table 2).

The stability of the ligand-CD58 complexes simulated in vacuum and in water was analyzed using RMSD profiles calculated over all atoms of the ligand-protein structures against their initial conformation (Online Resources 3 and 4, Figs. OR3.1 and OR4.1, respectively). As expected, the presence of the solvent destabilizes the complex and the profiles adopt larger RMSD values. Larger fluctuations are also observed in water due to reorientation of the ligand, as discussed later. Differences between RMSD values in vacuum and in water are apparently reduced for the larger ligand RTD-c where profiles coincide after about 30 ns during the production stage. It actually reflects a partial overlap of the ligand poses at the CD58 surface, while ligands P6 and P7 clearly adopt distinct poses at the CD58 surface depending upon the simulation conditions (Fig. 3 and Online Resource 3, Fig. OR3.4).

Ligand flexibility. The solvated peptide ligands undergo few conformational changes as confirmed by a cluster analysis, carried out with a RMSD cutoff of 0.15 nm (Table 2). Conformations of the solvated ligands P6 and P7 are gathered in only one cluster while RTD-

c is more flexible, with two conformation clusters found. Nevertheless, in that last case, over 99.99 % of the saved conformations belong to a single cluster. An analysis of the intramolecular φ-ψ angles shows that only local parts of the ligands are affected by conformational changes, especially located at the level of the turn inducers of P7 and RTD-c. Despite its size, the RTD-c conformation variety is strongly limited due to the presence of disulfide bridges. P6 also presents a very limited flexibility (Online Resource 6).

Ligand mobility. The location of the ligands on the CD58 surface is rather constrained and no long term Brownian diffusion regime is observed. In water, all three ligands occupy a different and limited region of space at the surface of the receptor (Fig. 3). Particularly, Tyr86 of the ligands appear to face, most of the time, the receptor surface. However, it does not form frequent Hbonds with CD58, with average values below 1 (Table 2). Only P7 presents an occurrence frequency of about 19.5 % for the Tyr86-Val17 Hbond (Table 1). Nevertheless, Tyr86 is involved in contacts with various CD58 residues, i.e., Leu38 and Ala45 for P6, Asn20 to Pro22, Leu38, Lys58, and Asn81 for P7, and Glu78 to Pro80, and Ile82 to Ser85 for RTD-c (Table 1). RTD-c thus binds to residues of CD58 that were identified using PDBePISA [19]. The translational mobility of P6 is more restricted than P7 and RTD-c, as illustrated by the short-term Mean Square Displacement (MSD) of the ligands calculated after alignment of the MD conformations onto the starting structure (Fig. 4a). P6 and RTD-c adopt a similar translational mobility trend, which is clearly slower than P7. Despite the lack of Brownian motion, tentative self-diffusion coefficient values D are obtained from the slope of the MSD curves (from 5 to 20 ns), divided by 6, i.e., 4.5 10^{-7}, 2.2 10^{-6}, and 7.1 10^{-7} nm^2/ps. On a reorientational point of view, P6 presents the slowest motion at very short term, as illustrated by the correlation function $<\cos\omega(t)>$ where ω(t) is the time-dependent angle variation of a vector perpendicular to three atoms of the ligand skeleton (Fig. 4b and Online Resource 7). Long-term τ_1 and short-term τ_2 correlation times, obtained from a two-exponential fit from 0 to 1 ns, are reported in Table 2. All parameters of the function fitting are given in Online Resource 7. The ligand P6 is characterized by the largest correlation time τ_1 = 37.644 ns associated with the slowest decrease of the long-term correlation function, while P7 and RTD-c adopt a similar reorientational behavior. P7 presents the smallest correlation time τ_2 = 0.0036 ns, associated with the fastest decrease of the correlation function, i.e., the highest short-term mobility.

Ligand-CD58 hydrogen bonds. In water, P6 and P7 form similar numbers of Hbonds with the solvent, i.e., 32 ± 3 and 29 ± 3, respectively, due to their similar size and amino acid

nature, while the larger RTD-c ligand forms an average of 61 ± 4 Hbonds (Table 2). Contrarily, the number of Hbonds formed between the ligands and CD58 is relatively similar between the three ligands, from 4 to 5. Particularly, the turn inducer in P6 and RTD-c form only very occasionally a single Hbond with CD58, while DBF of P7 more often involves one Hbond with Lys58 of CD58 (14.7 %) through its O atom. A detailed analysis of the intermolecular Hbonds numbers and occurrence frequency values are reported in Fig. 5 and Table 1, respectively. The Hbond occurrence is rather stable for P6 and RTD-c, while P7 produces a profile in agreement with the ligand-CD58 short-range interaction energy profile (Fig. 6), as detailed later using correlation indices κ. As already stated above, the ligands are not placed at the expected crystallographic binding site. Thus, Hbonds determined from the MD simulations differ from the crystallographic ones (Table 1). As an example, Tyr86 of the ligands form unfrequent Hbonds. In P6, the tyrosine residue is mostly placed in a parallel fashion from the CD58 surface as illustrated for the last frame generated during the 300 ns MD simulation (Fig. 7). It explains the low number of Hbonds formed during the whole simulation (Table 2). In P7, Tyr86 is Hbonded to Val17 of CD58 with a frequency of 19.5 %, and such Hbonding actually occurs during a limited period of time, i.e., between 13.516 and 25.542 ns. In the crystal structure, Tyr86 is preferentially Hbonded to Lys29. Residue Asp32, which interacts with Lys34 in the crystal structure, is only Hbonded in the solvated P6 complex, to Arg44 with a very large occurrence frequency value of 92.4 %, expressing the low mobility of the ligand. In P7, Asp31 is Hbonded to Thr57 with a relatively low frequency of 28.8 %, while it is expected to bind to Arg44 in the crystal structure. Lastly, the largest ligand RTD-c presents the highest number of Hbonds types, with occurrence frequencies close to or below 50 %.

Ligand-CD58 interaction energy. Even when the total energy of the system has converged, the ligand-CD58 short range interaction energy may vary, e.g., with the presence of Hbonds, or depending upon the orientation of the ligand facing the CD58 surface. The intermolecular interaction energy profiles displayed in Fig. 6 for the solvated ligand-CD58 complexes show that, for P6 and RTD-c, the energy values largely fluctuates around a mean energy value, while it adopts an oscillating behavior for P7. Based on the absolute ligand-CD58 intermolecular interaction energy, divided their solvent accessible surface (SAS), ligand efficiency indices were evaluated, leading to values around 38, 34, and 26 for P6, P7, and RTD-c, respectively (Table 2). The ranking of the ligands based on their efficiency index

is in agreement with the experimental IC_{50} value, i.e., 6.9 nM, 11.1, and 27 nM for P6, P7, and RTD-c, respectively.

As mentioned above, in TIP3P/SIRAH water, there is a rather important negative correlation between the short-range ligand-CD58 intermolecular interaction energy (Fig. 6) and the number of ligand-CD58 intermolecular Hbonds (Fig. 5). For instance, a value of -0.857 is obtained for ligand P7 (Table 2). Correlation factors are calculated using:

$$\kappa = \left[\frac{\sum_{i=1}^{N} (Hb_i . E_i)}{N} - \overline{(Hb.E)} \right] \frac{1}{\sigma_{Hb} \sigma_E} \quad (1)$$

where *Hb* and *E* stand for the number of ligand-CD58 Hbonds and the ligand-CD58 short-range interaction energy values, respectively. It suggests that the formation of Hbonds between the ligands and the receptor efficiently stabilizes the ligand facing the receptor when the system is solvated. Indeed, such a correlation is less clear in vacuum (Table OR3.1). In the case of solvated P7, one observes that, e.g., at t = 11.254 and 20.420 ns, the ligand forms 0 and 9 Hbonds with the receptor, respectively. Corresponding P7-CD58 interaction energy values are equal to -13.8 and -719.4 kJ/mol. In the first case, the distance between Tyr86 of the ligand and CD58 is equal to 0.734 nm while it is equal to 0.175 nm in the last case where it is Hbonded to three CD58 residues, i.e., Val17, Ser19, and Ser59. The remaining Hbonds involve several other residues of P7, i.e., Ile85 and Lys34 Hbonded to Asn20, Asp87 and all three aspartate residues, while the DBF moiety is pointing away from the complex structure (Fig.7).

Water-mediated ligand-cd58 hydrogen bonds. Due to their size, no SIRAH water coarse grains are located at distances below 0.35 nm from both the ligands and CD58, and the ligand-CD58 interface contains all-atom TIP3P water molecules only. Radial distributions functions (RDF) of the interface water molecules facing the ligand and protein surfaces show that, compared to their O atom, their H atoms can be located closer to, and consistently farther from, both the ligand and CD58 (Fig. 8). The g(O-surface atoms) RDFs let appear a small peak, around 0.2 nm, attributed to the interactions between the O atoms of the water molecules with Hbond acceptors of the ligand or CD58, followed by a larger one, centered around 0.26-0.27 nm, due to the interactions with heavy atoms of the ligand or CD58 [39,40]. In addition, the ligand can be indirectly bound to CD58 through bridging water molecules. To analyze such water-mediated Hbonds, water molecules that surround both the ligand and

CD58 within a distance of 0.35 nm were first identified every 1 ns during the production stage. Then, the water molecules that are Hbonded to both the ligand and CD58 were determined. As shown in Online Resource 8, P7 and RTD-c involve up to seven bridging water-mediated Hbonds, and only few RTD-c conformations actually present no bridging water molecules at all. Interestingly, P6 and RTD-c are characterized by the lowest mean number of bridging Hbonds, i.e., 1.3 (Table 2), and P6 has a maximum of only five of such water-mediated Hbonds. Selected conformations with a high number of bridging Hbonds are illustrated in Fig. 9. As examples, at t = 61 ns, P6 forms five bridging Hbonds, i.e., Tyr86-Glu42, Asp32-Glu37, Asp87-Ala45 (twice), and DPro-Ser47. At t = 210 ns, P7 forms three bridging Hbonds involving Asp31-Ser1, Asp31-Thr57, and Asp32-Leu23. For RTD-c, at t = 195 ns, seven bridging Hbonds are found, i.e., Asp31-Lys32, Glu36-Glu25, Ala18-Glu76 (three times), Tyr86-Glu78, and Asp87-Ser85.

Bridging water molecules are rather mobile. Their residence time was estimated from the average number of water molecules continuously acting as bridging molecules as a function of time (Fig. 10). On the average, only ten percents of the initial number of bridging H_2O remain after 0.5, 0.7, and 1 ns, for P6, P7, and RTD-c, respectively. Those figures are consistent with the small size of the ligands, especially for P6 and P7. The fit of a two-exponential function, from 0 to 5.5 ns allows to evaluate long-term and short-term residence times t_{r1} and t_{r2}, respectively (Table 2 and Online Resource 7). It appears that a mixture of slow and fast water molecules are involved as bridging partners between the ligand and CD58. In particular, the residence time of the fastest molecules, t_{r2}, follows the same trend as the IC_{50} values. A more detailed analysis of the P6-CD58 complex shows that among the 300 selected frames, only four water molecule appear in three consecutive frames, consistently with the short residence time. Thus, most of the molecules bridged to P6 are characterized by a fast exchange. Contrarily, for P7, two water molecules behave very differently. One of them has an occurrence number of 52 (over 300). It actually bridges Asp32 to Leu23 of CD58 during 52 ns, while the other bridges Lys34 to Lys58 of CD58 with a residence time of 15 ns. Regarding RTD-c, the highest residence time, 11 ns, is observed for a water molecule that links Tyr86 to Glu78 of CD58 (Fig. 9), while Lys34 is bridged to Pro80 through two water molecules acting separately, each with a residence time of 8 and 9 ns. The presence of long-living water molecules associated with P7 and RTD-c is in agreement with their larger values of t_{r1}. As P7 is the most mobile ligand, it is concluded that its mobility is not refrained by the long-living water-mediated Hbonds. Rather, the mobility mainly depends on direct Hbonds.

Several amino acid residues of the three ligands never or only rarely form direct Hbonds with CD58, such as the isoleucine residues in P6 and P7, the proline residues in P6, and Cys3, Cys5, and Cys7 residues in RTD-c (Table 3). A detailed analysis follows for the other ligand residues contained in CD2. Among the three ligands, P6 involves residue Asp32 which forms the highest number of direct Hbonds with CD58, i.e., 2.6 ± 0.7, mainly with Arg44 (Table 1), but also highly frequent bridging Hbonds with Glu37 only (Table 3). Such larger numbers of Hbonds and their corresponding occurrence frequencies can explain the lowest mobility of P6. DPro interacts with CD58 mainly through bridging Hbonds only. Asp87 also frequently interact through bridging Hbonds, with 107 occurrences, mainly with the Ala45 backbone (Table 3). It forms direct Hbonds as well, with Ser47 and, to a lesser extent, with Lys50 (Table 1). Residue Lys34 of P6 never forms any bridging Hbond, contrarily to P7 and RTD-c. Indeed, Lys34 of RTD-c involves 132 bridging Hbonds over the 300 selected frames of the production stage, most of them, 114, with the O atom of Pro80 of CD58 (Table 3). For P7, all residues but Ile85 form bridging Hbonds with CD58. It is particularly verified for Asp31, with 170 occurrences over the 300 analyzed frames, contrarily to P6 and RTD-c.

To summarize, MD frames provide ligand poses that differ from the C and F β-strands of CD2 in interaction with CD58, with different CD58 residues playing a role in the ligand-CD58 interactions. A reasonable correlation is observed between the number of ligand-CD58 Hbonds and the ligand-CD58 intermolecular interaction energy values in water. Thus, the experimentally observed transient binding character of a ligand towards CD58 can be due to the presence of non-permanent bridging water molecules that favor the mobility of the ligand facing the receptor surface. The experimentally larger IC_{50} value of P6 is correlated with the relatively larger number of Hbonds it can form with CD58, its low mobility and flexibility within the complex structure, and its larger efficiency index. It is also interesting to note the negative correlation between the IC_{50} values and the residence time of the fastest bridging water molecules, t_{r2}. Finally, P6 presents a reduced number of contact atoms facing CD58 compared to P7 and RTD-c, but this descriptor does not allow to differentiate P7 and RTD-c.

Molecular docking calculations

In a parallel work, ligand-protein molecular docking calculations were performed with constraints regarding the location of the ligands facing the CD58 structure. The aim was to evaluate a fast way to get optimal, and possibly multiple positions, of such peptides without

performing long MD calculations. Protein-protein docking programs were also considered for testing purposes. The obtained results are compared with X-ray structural information.

To evaluate the agreement between the ligand poses of the 20 ligand-protein docking calculations and the corresponding residues of the CD2 crystal structure, RMSD values were first calculated between the amino acid atoms of the ligand poses and their corresponding positions reported in the crystallographic CD2 structure (Fig. 2). The RMSD of the first 10 poses were determined and the minimal value of each docking is reported in Online Resource 9. It shows that no poses adopt a total RMSD value lower than 3.14 Å. Gold provides such a low RMSD value for the reference ligand REF only. Once larger ligands, with a non-crystal structure, are studied, RMSD values are all beyond 6 Å, even for the first best poses (Online Resource 10). It is explained by the fact that only a limited amount of interactions can be formed between CD58 and the ligands, while the whole CD2 structure forms a more complex and extended anchorage network (Online Resource 11). Indeed, a detailed analysis of the interface using PDBePISA reports that 18 and 21 residues are involved in the interface, respectively for CD58 and CD2. Additionally, the MD calculations reported above have shown that the ligands are slightly displaced from the binding site defined in the CD2-CD58 crystal structure. Online Resource 10 reports that a single docking calculation, i.e., Gold with *m_CFstrands* and a radius of 15 Å, provides a first best solution which is also the closest one to the experimental pose, for ligand REF with a RMSD of 4.69 Å. Thus, rather than focusing on the pose of the whole ligand structure, all possible ligand-CD58 distances were calculated between all residue side chains and CD58. It was achieved for the first three best poses of all 20 docking computations of the four ligands. For each ligand, the clustering of those 60 poses was achieved with the command *g_cluster* in Gromacs4.5.5. Using a RMSD cutoff value of 1.5 Å, i.e., the same value as used to cluster MD trajectories, 35, 51, 42, and 40 clusters were obtained, with a maximal size of 10, 2, 5, and 9 poses, for REF, P6, P7, and RTD-c, respectively. As expected, a larger diversity of solutions is obtained than in MD calculations. An analysis of the shortest contacts observed below 3.5 Å as reported in Online Resource 12 is discussed below. The very best pose in terms of score value, obtained with Autodock, Smina, and Gold, are displayed in Online Resource 13.

Reference ligand. For ligand REF, Gold appears to provide a very best solution that is close to the reference C and F β-strands of CD2 (Online Resource 13) but Tyr86 is located in a different binding site (Online Resource 14b). Indeed, in the crystal structure, Tyr86 is close to Lys34 and Phe46, with Lys34 being sandwiched between the two aromatic residues [9],

while in the Gold pose, Tyr86 faces Lys34 and Phe46 at distances larger than 3.5 Å in a different arrangement (Fig. 11b). Distance values are detailed in Table 4.

The very best poses of Vina and Vinardo are mutually similar (Fig. 11a and Online Resource 14a), but it is not systematically observed for all ligands. In both poses, Tyr86 faces Glu25 and Leu38 of CD58. Due to the receptor and ligand flexibility and the presence of all H atoms, Gold generates solutions that are characterized by shorter contact distances than Autodock and Smina. Residue Phe46 of CD58 is also more accessible and occurs at very short distances, below 2 Å, from Lys34 and Tyr86 only, while it forms no contacts when using Autodock and Smina. Distance values reported in Online Resource 12 can be analyzed so as to let clearly appear the most likely ligand locations and orientations after application of three pruning rules, i.e., (i) the discard of the ligand residues characterized by at least three docking calculations providing no contacts below 3.5 Å, (ii) the consideration of contact distances shorter than 2.1 Å only, and (iii) among the remaining contacts, the selection of the ones with a high occurrence number, above 5, considering all docking calculations (Table 5). Seven frequently observed contacts are left, involving Asp31, Asp32, Lys34, and Tyr86 of the peptide ligand. The three rules mentioned above are applied below to discriminate between the poses of the ligands P6, P7, and RTD-c using the set of ligand-protein molecular docking results.

Ligand P6. As the geometry of the P6 residues differ from the corresponding crystallographic structure, the poses obtained for ligand P6 do not let clearly appear expected contacts. Online Resources 13 and 14c show that the very best poses generated by Vinardo and Gold are almost identical. In the Vinardo best pose, the side chain of Tyr86 is packed against the hydrophobic part of the side chain of Lys34, as it is in the crystal state (Fig. 11c and Table 4). Tyr86 and Asp32 interact with Lys34 and Lys29 of CD58, respectively. Particularly, a Hbond is found by Gold between Asp32(OD1) and Lys29(HZ1). Rather, in their docking studies, Gokhale et al. [10] report two Hbonds occurring between Asp31 and Lys34, as well as an interaction between Tyr86 and Phe46, which do not correspond to the crystallographic orientation, but are among contacts also found in our calculations (Online Resource 12). The application of the three rules defined previously for REF, to the data in Online Resource 12, provide results reported in Table 5. It is seen that six frequent contacts are left with, among them, two experimentally observed ones, i.e., Asp32-Lys34 and Tyr86-Lys34. As for REF, the most frequent contacts appear between charged residues and/or tyrosine, consistently with the electrostatic properties of the CD58 surface.

Ligand P7. Ligand P7, the shortest one, is characterized by a large planar non-amino acid β-turn inducer, the DBF moeity. It very closely interacts with positive Lys32, Lys34, Arg44, or with negative Asp33, Glu78, Glu39, as well as with neutral Phe46 (Online Resource 12), while in MD, it was found to interact with CD58 mainly through bridging Hbonds. Many contacts with CD58 thus occur, contrarily to the β-turn inducers of P6 and RTD-c. Similarly to DBF, Lys34 of P7 presents numerous contacts with CD58, as also observed for ligands REF and P6. Regarding Tyr86, the very best poses generated by Autodock and Vina both show the residue in the same site (Fig. 11d and Online Resources 13 and 14d). The side chain is located at the basis defined by the four side chains of the charged Glu37, Glu39, Glu42, and Arg44 residues (Fig. 11d and Table 4). Finally, the application of the three rules mentioned previously provides a selection of contact sites reported in Table 5. Among the four most frequent contacts, only one, i.e., the contact involving Tyr86, was already involved by REF and P6.

Ligand RTD-c. RTD-c is the largest ligand built by an alternation of CD2 residues and disulfide bridges occurring between the two β-strands. Residues Lys34, Glu36, Lys82, and Tyr86 present numerous contacts with a variety of CD58 amino acid residues. Among them, only Tyr86 is frequently in contact with the expected residues Lys29, Lys32, and Lys34 (Online Resource 12). Indeed, Online Resource 14e and Fig. 11e show that the best pose generated by Gold presents Tyr86 in the same orientation as in the crystal structure, i.e., in interaction with Lys29, Lys32, and Lys34 of CD58. More precisely, Tyr86 is packed against the hydrophobic part of Lys34 of CD58 [9] (Table 4). The cysteine residues, particularly Cys7, Cys12, and Cys16, can generate close contacts with CD58, with positively-charged residues like Lys29, Lys34, Lys34 and Arg44, respectively. Cys16-Lys34 is one of the most frequent contacts (Table 5), and is obtained with Autodock and Smina (Online Resource 15a). The figure illustrates the 12 poses that present such a contact. Nine of them belong to the same cluster identified through the clustering of 60 poses described earlier in the text, and it is assumed that they reinforce the weight of the obtained pose. The results are thus consistent with the observation made from MD simulations, which showed that Cys16 is constantly forming a Hbond, with either Lys29 or Lys32 (Online Resource 3, Fig. OR3.7). Contacts between Cys16 and Arg44 are also detected when using Gold (Online Resource 15b). The figure actually reports four extremely similar poses. Finally, the application of the three rules mentioned previously provides a selection of contact sites reported in Table 5.

Only three most frequent contacts are left, involving the experimentally identified hot spots Tyr86 and Lys34.

To summarize, in ligands REF and P6, only residues Asp32 and Tyr86 are found to be close to their corresponding location in the crystal structure of CD2. For all four ligands, Tyr86 is always in contact with Lys34 even at distances shorter than the experimental ones. Most of the ligand residues form a close contact with the expected CD58 residue of the crystal structure, but not simultaneously. Lys34 of CD58 is often a binding spot for the ligands, regardless whether it is considered as the center of the search space or not. Lys34 is thus assumed to be a strong anchor point for a ligand. According to Table 5, conformations of the residues are important, but also their sequence. There are less and less highly frequent contacts that are recurrent between the 20 docking approaches used in the present work. Among them, there are less and less expected contacts. With respect to ligand REF, P6 has the most relevant contacts. Importantly, the insertion of structural constraints introduces a risk to miss expected contacts, as in ligands P7 and RTC-D. It is thus consistent with the trend adopted by the experimental IC_{50} values.

The ligand poses obtained from protein-protein docking calculations carried out using various web servers, i.e., ZDOCK [33], FRODOCK [34], ClusPro [35], and HADDOCK [36] are reported in the Online Resource 16. As for the ligand-docking programs, the search space must be constrained to allow ligands to stay close to the experimental binding site of CD58. It was achieved with ZDOCK and HADDOCK. In such cases, the binding site was defined using the 18 amino acid residues of CD58 that were identified using PDBePISA [19]. A variety of solutions were obtained, similarly to the ligand-protein docking programs. Also, some of the programs did not allow the proper treatment of DBF, the non-amino acid group of P7, like ClusPro and HADDOCK.

Conclusions and perspectives

Three cyclic ligands, named P6, P7, and RTD-c, forming complexes with CD58 were studied using Molecular Dynamics (MD) simulations and molecular docking calculations. These cyclic ligands, constrained through various β-turn inducers, are built so as to mimic the C and F β-strands of the protein CD2 in interaction with the receptor CD58, as described in literature. Ligand P6 is the closest one to the residue sequence of CD2. Results were analyzed

to emphasize most recurrent ligand-receptor contacts and interactions, particularly direct and bridging hydrogen bonds (Hbond) as well as short contacts, and were compared to the CD2-CD58 experimental structure (PDB ID 1QA9). In the crystal structure, the aliphatic branch of Lys34 in CD58 is packed between the aromatic rings of Phe46 and Tyr86 of CD2 [9].

MD simulations were carried out in vacuum and in a hybrid TIP3P/SIRAH all-atom/coarse grained water media. In vacuum, the ligand conformations remain rather constrained during the simulations, but their locations facing CD58 slightly differ from the CD2-CD58 crystal structure. In water, the ligands are more mobile. Particularly, P7 is placed at a location that strongly differs from the experimental binding area of CD58. The number of direct ligand-CD58 Hbonds correlates well with the ligand-CD58 interaction energy. The ligands are located farther away from the CD58 surface and interact with the receptor surface through bridging water molecules. For all ligands, water-mediated bridging Hbonds involve Asp32, Asp87, and to a variable extent, Tyr86. For P6, direct Hbonds are reinforced in water and Lys34 never forms any bridging Hbond, contrarily to P7 and RTD-c. Finally, P6 is the less mobile ligand at the CD58 surface. No particular influence of the turn inducer was found regarding their interaction with the receptor and the solvent. They all show a reduced amount of direct and/or bridging Hbond interactions with the receptor, except for DBF in P7 which interacts with many different CD58 residues through water-mediated bridging Hbonds.

The experimentally larger IC_{50} value of P6 is correlated with its relatively larger number of Hbonds formed with CD58, its low mobility and flexibility within the complex structure, and its large efficiency index. The index is calculated as the ratio of the short-range intermolecular ligand-CD58 energy and the ligand solvent accessible surface. It is also interesting to note the negative correlation between the IC_{50} values and the residence time of the fastest bridging water molecules, t_{r2}. Finally, P6 presents a reduced number of contact atoms closely facing CD58 compared to P7 and RTD-c. The special location of P7 at the surface of CD58 suggests that several possible binding sites might lead to an inhibition activity of the ligands, and is not inconsistent with the reported transient character of the CD58 complexation.

The clustering of the vacuum MD conformations yielded starting points for the docking calculations. A cavity cannot be unambiguously defined as the CD2-CD58 interface is rather flat and is highly electrostatic in nature. Therefore, the experimental crystal data were used to limit the search space. Several definitions of the search space were considered through four docking approaches, i.e., Gold, Vina, Vinardo, and Autodock. Despite the variability in the

docking results, a consensus can be found which lets appear robust contact sites identified to be Tyr86 in the ligands and Lys34 in the receptor, as determined experimentally [9]. Among the three ligands studied, P6 presents the highest number of highly frequent contacts that are also identified experimentally. Tyr86 and Lys34 form essential hot spots, that cannot, however, be considered as sufficient to unambiguously define the CD58 binding site for the ligands. A detailed analysis of the very best scored pose identified in each of the molecular docking family, i.e., Autodock, Vina, Vinardo, and Gold, allows to characterize other interaction modes. An additional contact, that is also seen as important for the ligand binding, involves Asp32 of the ligands and Lys34 of CD58. It is also observed that Lys34 of CD58 can be in contact with a variety of ligand residues, such as Ile85 in P6, Asp31 and Lys34 in P7, and Cys16 in RTD-c. Finally, Asp33 of CD58 is involved in a number of contacts with the ligands, such as Tyr86 of P6 and Lys34 of P7. Several protein-protein molecular docking programs were also considered. Similarly to ligand-protein docking results, their use shows that the binding site of CD58 must be *a priori* known. Some of the tested programs also did not allow the proper treatment of the non-amino acid nature of the P7 β-turn inducer.

MD and molecular docking results are not stricly comparable as the later are being obtained using constrained search spaces. Additionally, several of the selected docking methods neglect most of the hydrogen atoms, and/or atomic charges, and are based on evaluation functions that strongly differ from any conventional force field. Under such conditions, molecular docking is forced to provide poses that involve residues playing a significant role within the CD2-CD58 crystal structure. MD results show the effect of an explicit solvent and let additionally appear the recurrent occurrence of Glu78 of CD58 in the ligand binding.

The three ligand-CD58 systems studied so far are now considered for testing an in-house protein electrostatic model built from Amber99SB atomic charges [41]. Further studies will also focus on the receptor deformation under binding with a ligand and a finer analysis, through QM/MM calculations, of the binding modes including direct and bridging hydrogen bonds. The central conformation obtained from the clustering analyses of the vacuum and solvated ligand-CD58 systems will be used as starting conformations of the QM/MM calculations. In a later stage, mutations in the ligand sequences will be studied as modulators of the Tyr86 role in the ligand binding.

References

1. Sable R, Jois J (2015) Surfing the protein-protein interaction surface using docking methods: Application to the design of PPI inhibitors. Molecules 20: 11569-11603
2. Zinzalla G, Thurston DE (2009) Targeting protein–protein interactions for therapeutic intervention: A challenge for the future. Future Med Chem 1:65-93
3. Wang JH, Smolyar A, Tan K, Liu JH, Kim M, Sun ZY, Wagner G, Reinherz EL (1999) Structure of a heterophilic adhesion complex between the human CD2 and CD58 (LFA-3) counterreceptors. Cell 97:791-803
4. Raychaudhuri S, Thomson BP, Remmers EF, Eyre S, Hinks A, Guiducci C, Catanese JJ, Xie G, Stahl EA, Chen R, Alfredsson L, Amos CI, Ardlie KG, BIRAC Consortium, Barton A, Bowes J, Burtt NP, Chang M, Coblyn J, Costenbader KH, Criswell LA, Crusius JB, Cui J, De Jager PL, Ding B, Emery P, Flynn E, Harrison P, Hocking LJ, Huizinga TW, Kastner DL, Ke X, Kurreeman FA, Lee AT, Liu X, Li Y, Martin P, Morgan AW, Padyukov L, Reid DM, Seielstad M, Seldin MF, Shadick NA, Steer S, Tak PP, Thomson W, van der Helm-van Mil AH, van der Horst-Bruinsma IE, Weinblatt ME, Wilson AG, Wolbink GJ, Wordsworth P, YEAR Consortium, Altshuler D, Karlson EW, Toes RE, de Vries N, Begovich AB, Siminovitch KA, Worthington J, Klareskog L, Gregersen PK, Daly MJ, Plenge RM (2009) Genetic variants at CD28, PRDM1 and CD2/CD58 are associated with rheumatoid arthritis risk, Nat Genet 41:1313-1320
5. Liu J, Li C, Ke S, Satyanarayanajois SD (2007) Structure-based rational design of β-hairpin peptides from discontinuous epitopes of cluster of differentiation 2 (CD2) protein to modulate cell adhesion interaction. J Med Chem 50:4038-4047
6. Gokhale A, Weldeghiorghis ThK, Taneja V, Satyanarayanajois D (2011) Conformationally constrained peptides from CD2 to modulate protein-protein interactions between CD2 and CD58. J Med Chem 54:5307-5319
7. Ikemizu S, Sparks LM, van der Merwe PA, Harlos K, Stuart DI, Jones EY, Davis SJ (1999) Crystal structure of the CD2-binding domain of CD58 (lymphocyte function-associated antigen 3) at 1.8-A resolution. Proc Natl Acad Sci U S A 96:4289-4294
8. van der Merwe PA, Barclay AN, Mason DW, Davies EA, Morgan BP, Tone M, Krishnam AK, Ianelli C, Davis SJ (1994) Human cell-adhesion molecule CD2 binds CD58 (LFA-3) with a very low affinity and an extremely fast dissociation rate but does not bind CD48 or CD59. Biochemistry 33:10149-10160

9. Kim M, Sun ZYJ, Byron O, Campbell G, Wagner G, Wang JH, Reinherz EL (2001) Molecular dissection of the CD2-CD58 counter-receptor interface identifies CD2 Tyr86 and CD58 Lys34 residues as the functional "hot spot". J Mol Biol 312:711-720
10. Gokhale A, Kanthala S, Latendresse J, Taneja V, Satyanarayana SD (2013) Immunosuppression by co-stimulatory molecules: Inhibition of CD2-CD48/CD58 interaction by peptides from CD2 to suppress progression of collagen-induced arthritis in mice. Chem Biol Drug Des 82:106-118
11. Gokhale AS, Sable R, Walker JD, McLaughlin L, Kousoulas KG, Satyanarayana SD (2015) Inhibition of cell adhesion and immune responses in the mouse model of collagen-induced arthritis with a peptidomimetic that blocks CD2-CD58 interface interactions. Biopolymers 104:733-742
12. Sable R, Durek T, Taneja V, Craik DJ, Pallerla S, Gauthier T, Jois S (2016) Constrained cyclic peptides as immunomodulatory inhibitors of the CD2:CD58 protein-protein interaction. ACS Chem Biol 11:2366-2374
13. Bayas MV, Schulten K, Leckband D (2003) Forced detachment of the CD2-CD58 complex. Biophys J 84: 2223-2233
14. Bayas MV, Kearney A, Avramovic A, van der Merwe PA, Leckband DE (2007) Impact of salt bridges on the equilibrium binding and adhesion of human CD2 and CD58. J Biol Chem 282: 5589-5596
15. Abdel-Azeim S, Chermak E, Vangone A, Oliva R, Cavallo L (2014) MDcons: Intermolecular contact maps as a tool to analyze the interface of protein complexes from molecular dynamics trajectories. Bioinformatics 15:S1
16. Wang X, Ji CG, Zhang ZH (2015) Glycosylation modulates human CD2-CD58 adhesion via conformational adjustment. J Phys Chem B 119, 6493-6501
17. Jining L, Makagiansar I, Yusuf-Makagiansar H, Chow VTK, Siahaan TJ, Jois SDS (2004) Design, structure and biological activity of β-turn peptides of CD2 protein for inhibition of T-cell adhesion. Eur J Biochem 271:2873-2886
18. Lawson ADG, MacCoss M, Heer, JP (2018) Importance of rigidity in designing small molecules drugs to tackle protein-protein interactions (PPIs) through stabilization of desired conformers. J Med Chem 61:4283-4289
19. PDBePISA Proteins, Interfaces, Structures and Assemblies v.1.52. http://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver. Accessed 26 Feb 2018

20. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M (1977) The Protein Data Bank: A computer-based archival file for macromolecular structures. J Mol Biol 112:535-542
21. Pietal MJ, Tuszynska I, Bujnicki JM (2007) PROTMAP2D: Visualization, comparison and analysis of 2D maps of protein structure. Bioinformatics 23:1429-1430
22. PyMol™ Molecular Graphics System, v.1.8.6.0, Schrödinger LLC, 2013
23. Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. J Chem Theory Comput 4:435-447
24. Pronk S, Páll S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B, Lindahl E (2013) GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 29:845-854
25. Showalter SA, Brüschweiler R (2007) Validation of Molecular Dynamics simulations of biomolecules using NMR spin relaxation as benchmarks: Application to the AMBER99SB force field. J Chem Theory Comput 3:961-975
26. Darré L, Tek A, Baaden M, Pantano S (2012) Mixing atomistic and coarse grain solvation models for MD simulations: Let WT4 handle the bulk. J Chem Theory Comput 8:3880-3894
27. Gonzales HC, Darré L, Pantano S (2013) Transferable mixing of atomistic and coarse-grained water models. J Phys Chem B 117:14438-14448
28. Darré L, Machado MR, Brandner AF, González HC, Ferreira S, Pantano S (2015) SIRAH: A structurally unbiased coarse-grained force field for proteins with aqueous solvation and long-range electrostatics. J Chem Theory Comput 11:723-739
29. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) Autodock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 16:2785-2791
30. Koes DR, Baumgartner MP, Camacho CJ (2013) Lessons learned in empirical scoring with smina from the CSAR 2011 benchmarking exercise. J Chem Inf Model 53:1893-1904

31. Trott O, Olson AJ (2010) AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem 31:455-461
32. Jones G, Willett P, Glen RC, Leach AR, Taylor R (1997) Development and validation of a genetic algorithm for flexible docking. J Mol Biol 267:727-748
33. Pierce BG, Hourai Y, Weng Z (2011) Accelerating protein docking in ZDOCK using an advanced 3D convolution library. PLoS One 6:e24657
34. Ramírez-Aportela E, López-Blanco JR, Chacón P (2016) FRODOCK 2.0: Fast protein-protein docking server. Bioinformatics 32:2386-2388
35. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, Beglov D, Vajdab S (2017) The ClusPro web server for protein-protein docking. Nature Protocols 12:255-278
36. van Zundert GCP, Rodrigues JPGLM, Trellet M, Schmitz C, Kastritis PL, Karaca E, Melquiond ASJ, van Dijk M, de Vries SJ, Bonvin AMJJ (2016) The HADDOCK2.2 webserver: User-friendly integrative modeling of biomolecular complexes. J Mol Biol 428:720-725
37. Forli S, Huey R, Pique ME, Sanner MF, Goodsell DS, Olson AJ (2016) Computational protein-ligand docking and virtual drug screening with the AutoDock suite. Nat Protoc 11:905-919
38. Feinstein WP, Brylinski M (2015) Calculating an optimal box size for ligand docking and virtual screening against experimental and predicted binding pockets. J Cheminform 7:18
39. Bizzarri AR, Cannistraro S (2002) Molecular Dynamics of water at the protein-solvent interface. J Phys Chem B 106:6617-6633
40. Dastidar SG, Mukhopadhyay C (2003) Structure, dynamics, and energetics of water at the surface of a small globular protein: A molecular dynamics simulation. Phys Rev E 68:021921
41. Leherte L, Vercauteren DP (2014) Evaluation of Reduced Point Charge Models of Proteins Through Molecular Dynamics Simulations: Application to the Vps27 UIM-1 - Ubiquitin Complex. J Mol Graphics Model 47:44-61

Figure captions

Fig. 1 (Top) Cartoon representation of the CD2 protein. Residues 31 to 34 and 84 to 87 are highlighted in dark blue and red, respectively. (Bottom) Planar and perspective structure of the ligands REF, P6, P7, and RTD-c. Conserved CD2 residues are in bold and are labeled as in the PDB structure file 1QA9. Ligand skeletons and side chains are shown using thick and thin lines, respectively. The 3D structure of DBF is presented in Online Resource 1, Fig. OR1.1

Fig. 2 Superimposition of CD58 (gray surface) and residues of the C and F β-strands of CD2 (black lines). Residues of CD58 located at a distance shorter than 4 Å from the CD2 strands are colored. H atoms are not shown for clarity

Fig. 3 Superimposition of 100 frames of the ligand-CD58 complexes in TIP3P/SIRAH water obtained from the 300 ns NPT MD simulations at 300 K and 1 bar. Tyr86 of the ligands is shown with orange sticks

Fig. 4 (a) MSD profiles and (b) reorientation correlation function of the ligand-CD58 complexes in TIP3P/SIRAH water obtained from the 300 ns MD simulations at 300 K and 1 bar

Fig. 5 Ligand-CD58 (black) and ligand-water (red) hydrogen bond profiles obtained from the 300 ns MD simulations, at 300 K and 1 bar

Fig. 6 Ligand-CD58 short-range intermolecular interaction energy in TIP3P/SIRAH water obtained from the 300 ns MD simulations, at 300 K and 1 bar

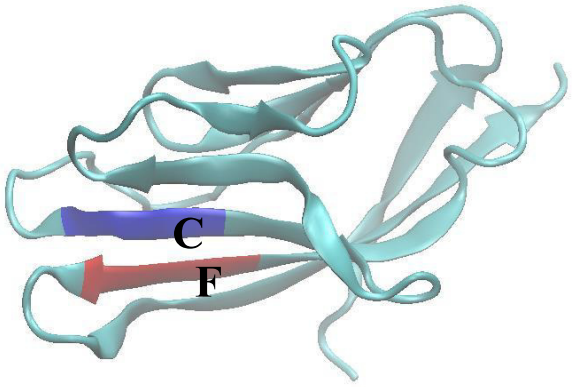
Fig. 7 Snapshots of the ligand-CD58 configurations at t = 300 ns for P6, at (a) t = 11.254 and (b) 20.420 ns for P7, and at 300 ns for RTD-c, in TIP3P/SIRAH water obtained from the 300 ns NPT MD simulations at 300 K and 1 bar. Tyr86 of the ligand is shown with orange sticks. Selected residues of CD58 are shown with green sticks (P6: Ala45 and Leu38, P7: Val17, Ser19, and Ser59, RTD-c: Glu78 to Pro80 and Ile82 to Ser85)

Fig. 8 Radial distribution functions of TIP3P and surface atoms of the ligand-CD58 complexes obtained from the 300 ns NPT MD simulations in TIP3P/SIRAH solvation conditions, at 300 K and 1 bar

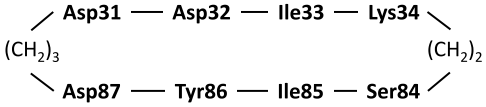
Fig. 9 Selected poses with bridging water molecules obtained from the 300 ns NPT MD simulations at 300 K and 1 bar. Water molecules located at a maximal distance of 0.35 nm from the ligand and the receptor are shown in gray. Among them, water molecules involved in water-mediated bridging Hbonds are displayed in green. Long-lasting bridging molecules are shown with purple sticks

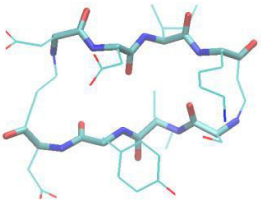
Fig. 10 Average number of water molecules continuously acting as bridging Hbond mediator as a function of time, obtained from the 300 ns NPT MD simulations in TIP3P/SIRAH solvation conditions, at 300 K and 1 bar.

Fig. 11 Superimpositions of very best poses in terms of score values of the molecular docking families (Online Resource 10). (a) Best poses (Vina: green; Vinardo: blue) of ligand REF. Glu25, Lys34, Leu38, and Phe46 of CD58 are colored in light blue, yellow, magenta, and gray; (b) Gold best pose (orange) of ligand REF. Lys34 and Phe46 of CD58 are colored in yellow and gray; (c) Vinardo best pose (blue) of P6. Lys34 and Phe46 of CD58 are colored yellow and gray; (d) Best poses (Autodock: red; Vina: green) of P7. Glu37, Glu39, Glu42, and Arg44 are colored in pale yellow, cyan, pink, and orange; (e) Gold best pose (orange) of RTD-c. Lys29, Lys32, Lys34, and Phe46 are colored in red, green, yellow, and gray. CD2 residues are in black. H atoms are not shown for clarity

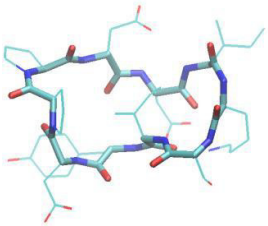


REF

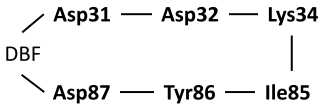


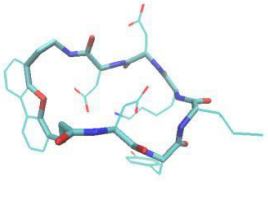


P6

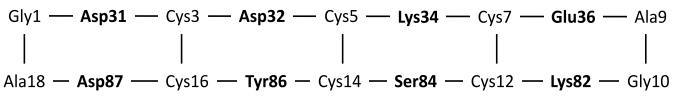


P7





RTD-c



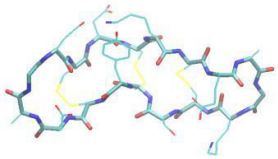


Fig. 1

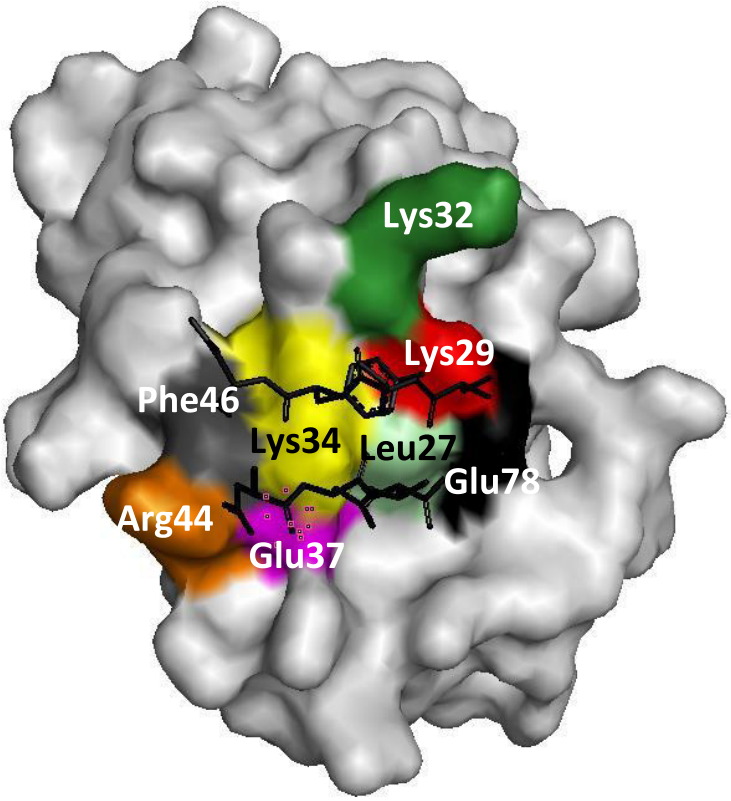
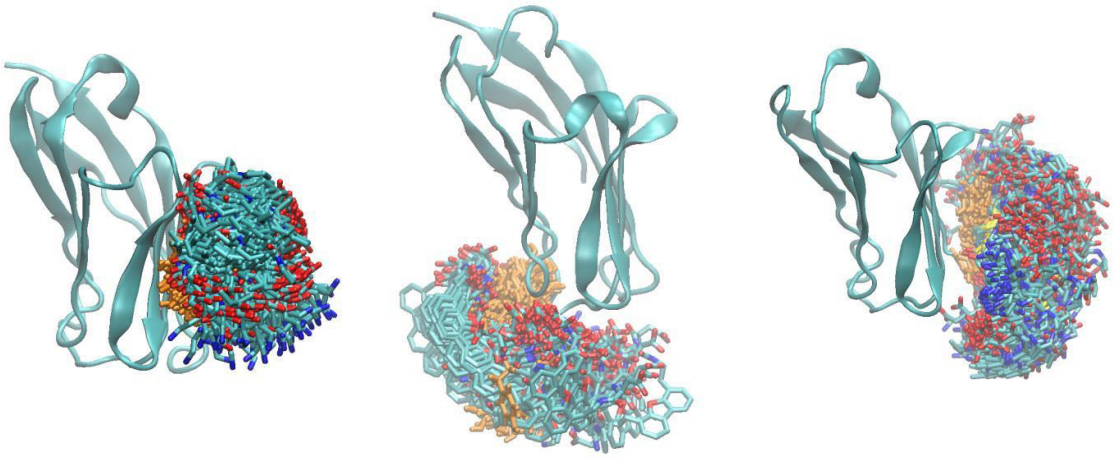
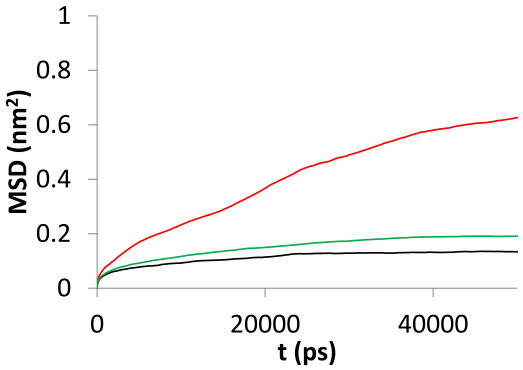


Fig. 2

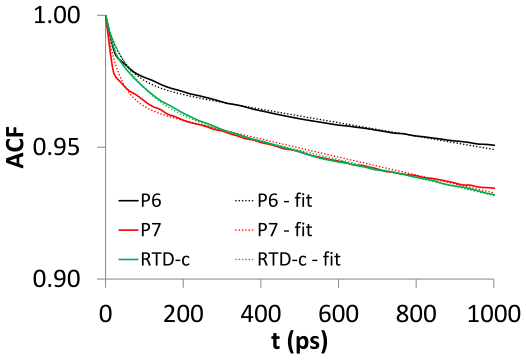


P6 P7 RTD-c

Fig. 3

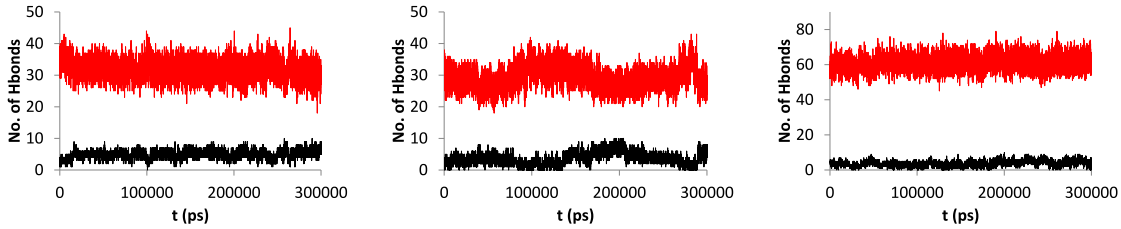


(a)



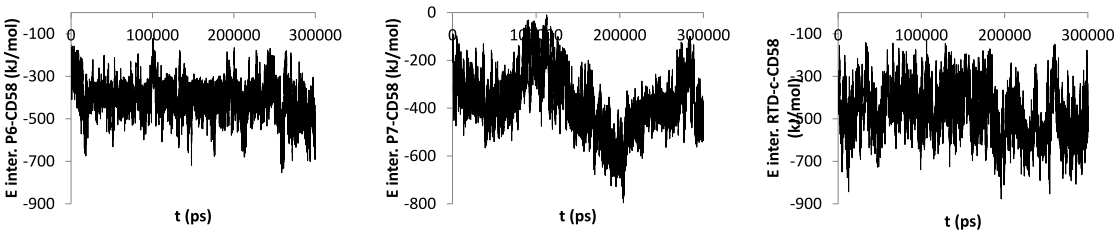
(b)

Fig. 4



P6 P7 RTD-c

Fig. 5



P6 P7 RTD-c

Fig. 6

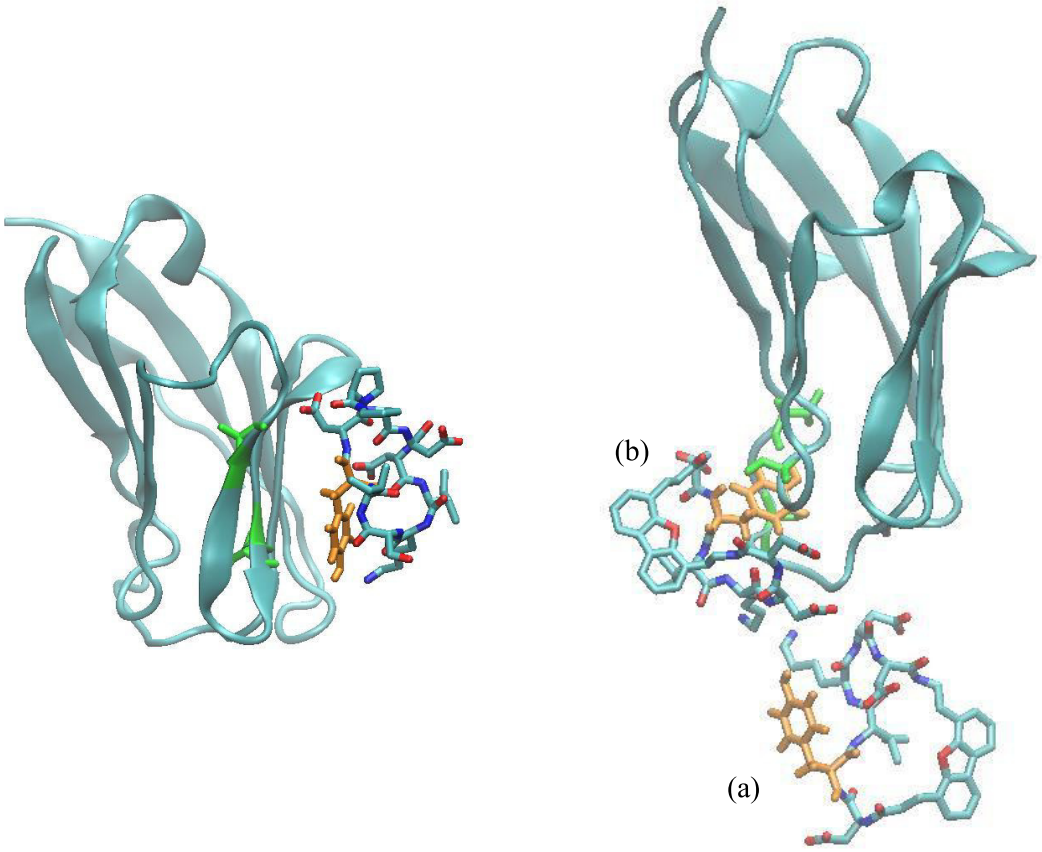


Fig. 7

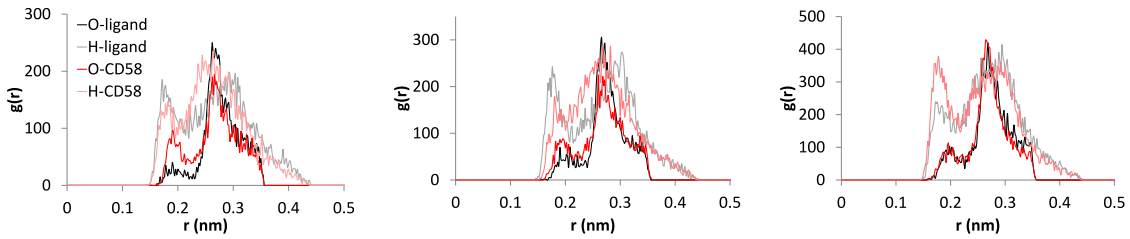
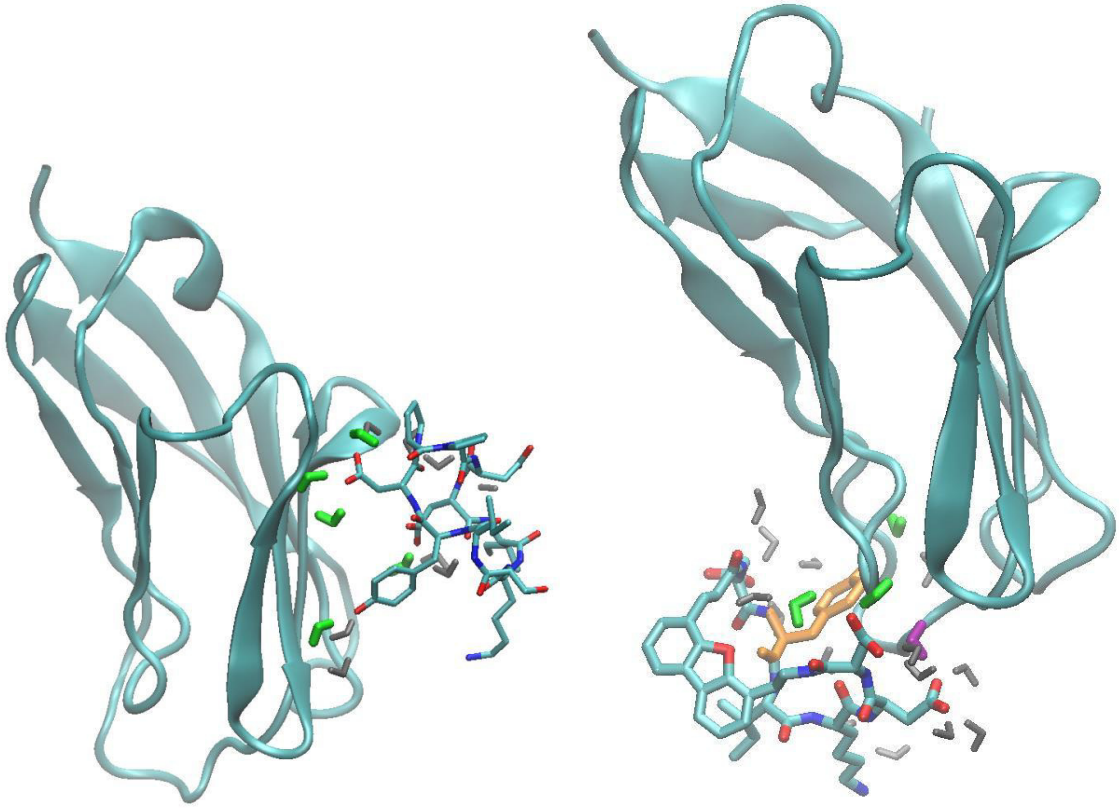
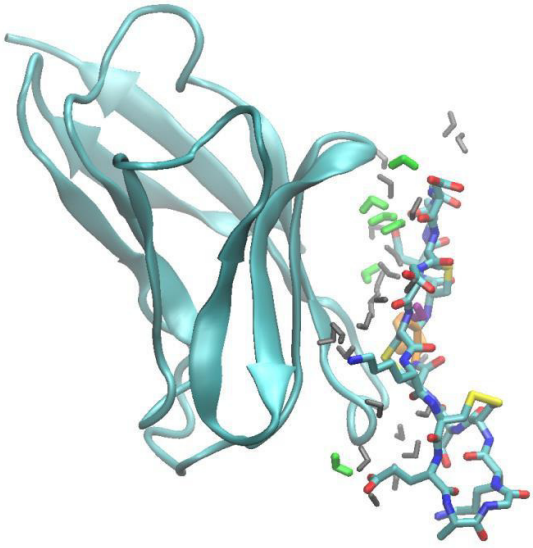


Fig. 8



P6 (t = 61 ns)

P7 (t = 210 ns)



RTD-c (t = 195 ns)

Fig. 9

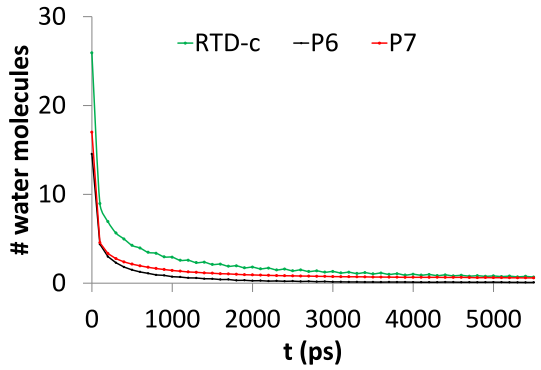
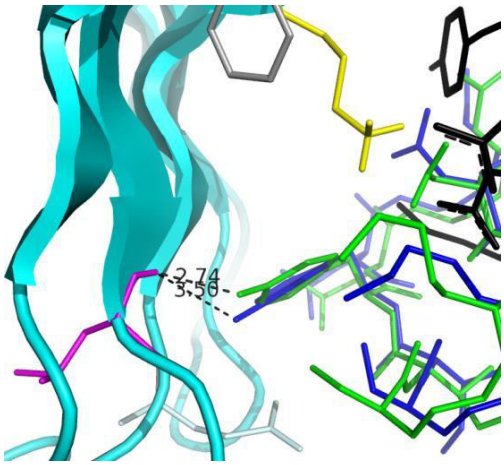
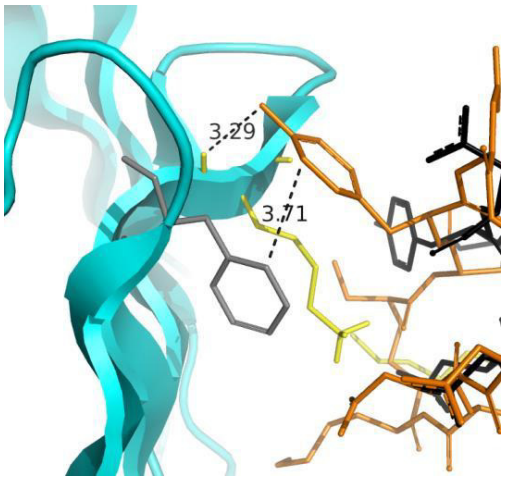


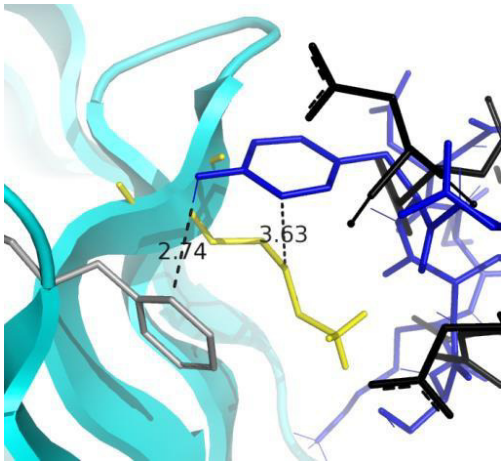
Fig. 10



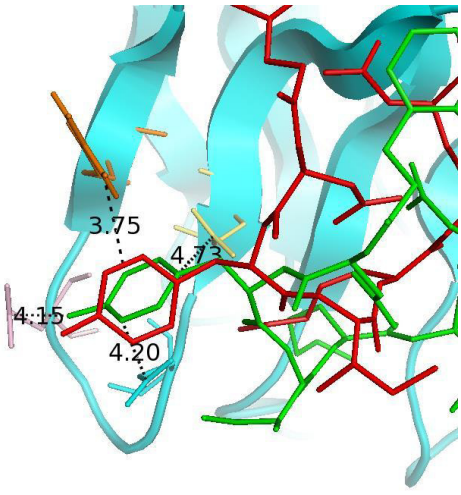
(a)



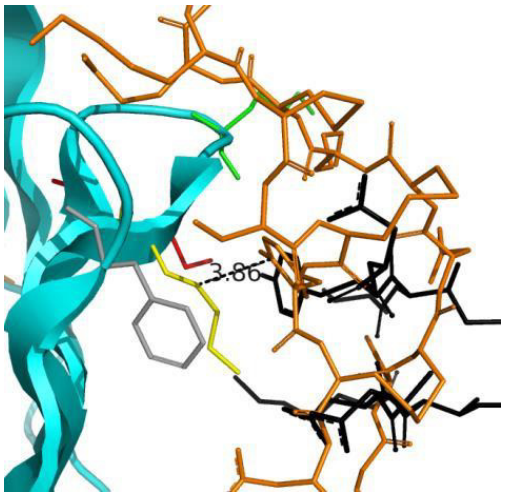
(b)



(c)



(d)



(e)

Fig. 11

Table 1 PDBePISA hydrogen bonds and Protmap2D contacts[a] between residues Asp31 to Lys34 and Ser84 to Asp87 of CD2 and the receptor CD58, as obtained from the crystallographic data file (PDB ID 1QA9). Ligand-CD58 hydrogen bonds[b] (occurrence frequency[c] in parentheses) and Tyr86-CD58 contacts[a] obtained from the 300 ns NPT MD simulations at 300 K and at 1 bar

		1QA9 crystal structure	MD simulations		
	CD58	CD2	P6	P7	RTD-c
Hbonds					
	Val17			Tyr86 (19.5)	
	Lys24			Asp32 (36.2)	Glu36 (50.5)
	Glu25				Lys34 (47.4)
	Lys29	Tyr86			
	Lys32				Asp87 (24.3)
	Lys34	Asp32			
	Glu39				Lys34 (27.8)
	Arg44	Asp31	Asp32 (92.4)		
	Ser47		Asp87 (67.3)		
	Lys50		Asp87 (21.4)		
	Thr57			Asp31 (28.8)	
	Thr83				Ser84 (37.2)
Contacts[d]					
	Asn20			x	
	Gln21			x	
	Pro22			x	
	Glu25				
	Leu27	Lys34			
	Lys29	Tyr86			
	Lys32	Tyr86			
	Asp33				
	Lys34	Asp32, Tyr86			
	Glu37	Asp32			
	Leu38		x	x	
	Glu39				
	Glu42				
	Arg44	Asp31			
	Ala45		x		
	Phe46	Asp31, Asp87			
	Glu47				
	Lys58			x	
	Glu78	Lys34, Glu36			x
	Ser79				x
	Pro80				x
	Asn81			x	
	Ile82				x
	Thr83				x
	Asp84				x
	Ser85				x

[a] Only contact distances shorter than 0.5 nm are reported

[b] Geometrical cutoffs selected to identify Hbonds were set to 0.35 nm and 30° for the acceptor-donor distance and the acceptor-donor-hydrogen angle, respectively

[c] Only occurrence frequency values larger or equal to 20 % are reported
[d] The 18 binding residues of CD58 identified using PDBePISA [19] are in bold

Table 2 Mean properties and standard deviations obtained from 300 ns MD simulations at 300 K and 1 bar

	P6	P7	RTD-c
Ligand-PDBePISA com-com dist. (nm)	1.360 ± 0.077	2.036 ± 0.105	1.235 ± 0.113
No. of ligand conformation clusters (RMSD = 0.15 nm)	1	1	2[b]
No. of Tyr86-CD58 Hbonds[a]	0.06 ± 0.28	0.57 ± 0.79	0.16 ± 0.38
No. of turn-CD58 Hbonds[a]	0.001 ± 0.032	0.3 ± 0.5	0.01 ± 0.10
No. of ligand-water Hbonds[a]	32 ± 3	29 ± 3	61 ± 4
No. of ligand-CD58 Hbonds[a]	5 ± 1	4 ± 2	4 ± 2
No. of bridging Hbonds[a]	1.3 ± 1.8	2.9 ± 1.0	1.3 ± 1.5
D (nm^2/ps)	$4.0\ 10^{-7}$	$2.1\ 10^{-6}$	$6.4\ 10^{-7}$
τ_1 (ns)	37.644 ± 1.102	27.684 ± 0.679	27.670 ± 0.668
τ_2 (ns)	0.050 ± 0.004	0.036 ± 0.003	0.086 ± 0.005
$\kappa\ E_{lig\text{-}CD58}/Hb_{lig\text{-}CD58}$	-0.711	-0.857	-0.793
t_{r1} (ns)	0.774 ± 0.036	2.812 ± 0.197	2.063 ± 0.108
t_{r2} (ns)	0.054 ± 0.002	0.059 ± 0.003	0.068 ± 0.004
Mean $E_{lig\text{-}CD58}$ (kJ/mol)	-406.26 ± 83.08	-372.54 ± 123.07	-473.51 ± 118.87
No. contact atoms ($d \leq 0.35$ nm)	24 ± 5	47 ± 14	44 ± 9
Ligand SAS (nm^2)	10.85 ± 0.34	11.05 ± 0.39	18.06 ± 0.49
Ligand efficiency	37.51 ± 8.00	33.67 ± 11.09	26.22 ± 6.52
Experimental IC_{50} (nM)	6.9 ± 0.4 [6,12]	11.1 ± 3.8 [6]	27 ± 15 [12]

[a] Geometrical cutoffs selected to identify Hbonds were set to 0.35 nm and 30° for the acceptor-donor distance and the acceptor-donor-hydrogen angle, respectively

[b] One cluster gathers > 99.99 % of the conformations.

Table 3 Occurrence of bridging water molecules[a] calculated over 300 frames during the production stage, and number of direct hydrogen bonds[b] formed between the ligands and CD58, as calculated from the 300 ns MD simulations at 300 K and 1 bar

	Ligand residue	Bridging water molecules	Bridged to	Direct H Bonds
P6	Ser84	0		0
	Ile85	0		0
	Tyr86	47	Glu37 (30), Glu39 (10), Glu42 (6)	0.1 ± 0.3
	Asp87	107	Ala45 (83), Ser47 (21)	1.9 ± 1.0
	DPro	23	Ser47 (23)	$\sim10^{-3}$
	Pro	0		0
	Asp31	0		$\sim10^{-4}$
	Asp32	218	Glu37 (218)	2.6 ± 0.7
	Ile33	0		0
	Lys34	0		0.2 ± 0.5
P7	Asp87	40	Asn20 (27)	0.3 ± 0.6
	Tyr86	33	Glu21 (9), Ser59 (14)	0.6 ± 0.8
	Ile85	0		0.2 ± 0.4
	Lys34	53	Gln21 (6), Pro22 (18), Lys58 (27)	0.3 ± 0.5
	Asp32	160	Gln21 (7), Pro22 (57), Leu23 (72), Ser41 (16)	0.5 ± 0.5
	Asp31	170	Pro22 (9), Leu23 (36), Asn40 (17), Ser41 (83), Thr57 (18)	0.7 ± 0.6
	DBF	70	Asn40 (8), Asp6 (30), Lys58 (24)	0.2 ± 0.4
RTD-c	Gly1	0		$\sim10^{-3}$
	Asp31	24	Cys32 (6), Glu76 (6), Glu78 (10)	0.1 ± 0.3
	Cys3	0		$\sim10^{-3}$
	Asp32	114	Glu25 (5), Lys32 (14), Glu37 (5), Glu78 (85)	0.2 ± 0.5
	Cys5	0		$\sim10^{-2}$
	Lys34	132	Glu78 (17), Pro80 (114)	1.1 ± 1.0
	Cys7	4		$\sim10^{-5}$
	Glu36	209	Glu25 (183), Pro80 (16), Asn81 (10)	0.8 ± 0.7
	Ala9	0		0
	Gly10	0		0
	Lys82	4		$\sim10^{-2}$
	Cys12	5		0
	Ser84	55	Pro80 (17), Asn81 (6), Ile82 (11), Thr83 (17)	0.6 ± 0.6
	Cys14	9	Thr83 (9)	$\sim10^{-2}$
	Tyr86	224	Glu76 (19), Met77 (10), Glu78 (161), Asp84 (20), Ser85 (8)	0.3 ± 0.5
	Cys16	1		$\sim10^{-4}$
	Asp87	93	Glu76 (69), Glu78 (5), Asp84 (5)	0.7 ± 0.8
	Ala18	11	Glu76 (10)	$\sim10^{-2}$

[a] Only values greater or equal to 5 are reported. [b] Geometrical cutoffs selected to identify Hbonds were set to 0.35 nm and 30° for the acceptor-donor distance and the acceptor-donor-hydrogen angle, respectively

Table 4. Atoms involved in the Tyr86-CD58 distance values that characterize the very best pose of the four molecular docking family, i.e., Autodock, Vina, Vinardo, and Gold, as illustrated in Fig. 11. Score values are given in Online Resource 10

	Docking method	Model	Tyr86 atom	CD58 atom	Distance (Å)	Fig. 11
REF	Vina	m_largesite	OH	Leu38 (HN)	2.74	a
	Vinardo	m_lys34	OH	Leu38 (HN)	3.50	a
	Gold	m_CFstrands	OH	Lys34 (O)	3.29	b
	Gold	m_CFstrands	CE2	Phe46 (CD2)	3.71	b
P6	Vinardo	m_smallsite	HH	Phe46 (CD2)	2.74	c
	Vinardo	m_smallsite	CE1	Lys34 (CD)	3.63	c
P7	Vina	m_smallsite	CG	Glu37(CD)	4.73	d
	Vina	m_smallsite	CE1	Glu39 (CD)	4.20	d
	Vina	m_smallsite	OH	Glu42 (CD)	4.15	d
	Vina	m_smallsite	CE2	Arg44 (NH1)	3.75	d
RTD-c	Gold	m_lys34 - R10	CZ	Lys34 (CG)	3.86	e

Table 5 Occurrence frequency of the shortest (< 2.1 Å) and most frequent (no. of occurrences > 5) ligand-CD58 contacts as observed from the first three poses of the docking calculations reported in Online Resource 12

REF		P6		P7		RTD-c	
Contact	Occ.	Contact	Occ.	Contact	Occ.	Contact	Occ.
Asp31 - Arg44	6	Asp32 – Lys29	9	Asp31 - Lys34	8		
Asp32 - Lys34	10	Asp32 – Lys34	7				
Lys34 - Glu37	12	Lys34 – Glu25	6	Lys34 – Lys34	10		
Lys34 - Glu39	11	Ile85 - Lys34	7	Lys34 – Asp33	6		
Tyr86 - Lys34	7	Tyr86 - Lys34	11	Tyr86 - Lys34	12	Tyr86 - Lys34	10
Tyr86 - Asp33	6	Tyr86 – Asp33	6			Asp87 - Arg44	7
Tyr86 - Phe46	6					Cys16 - Lys34	12